



فرائن للعلوم الاقتصادية والإدارية

KHAZAYIN OF ECONOMIC AND
ADMINISTRATIVE SCIENCES
ISSN: 2960-1363



Abayesian-Elasticnet Hybrid for Estimating the Gama Regression Model in the presence of Multicollinearity With Application

Gheadaa Ibrahim Shehab
Diyala University, Baqubah, Iraq
gh.gheadaa@uodiyala.edu.iq

Abstract. In this research, the Bayesian approach was employed to estimate the parameters of the Gamma regression model in the presence of the multicollinearity problem. As is well known, the Bayesian approach becomes inefficient in models suffering from multicollinearity, as the estimated parameters tend to be inefficient. Therefore, to enhance the performance of this approach, one of the penalization techniques, namely the Elastic Net technique, was utilized to estimate the parameters of the Gamma regression model under multicollinearity. The study aimed to investigate the behavior of the Bayesian approach in parameter estimation when this technique is applied, as well as to compare the proposed technique with classical methods. The results demonstrated the superiority and effectiveness of the employed technique in the estimation process. Simulation experiments were conducted in addition to real data analysis by computing the mean squared error at different levels of correlation among the independent variables and under varying sample sizes

Keywords: Abayesian-Elasticnet, Gama Regression, The problem of multicollinearity

DOI: 10.69938/Keas.2603031

في الاسلوب البيزي لتقدير انموذج انحدار كاما في ظل Elastic net تهجين تقنية وجود مشكلة التعدد الخطي مع التطبيق

غيداء ابراهيم شهاب
جامعة ديالى، بعقوبة، العراق
gh.gheadaa@uodiyala.edu.iq

في هذا البحث، استُخدم المنهج البايزي لتقدير معالم نموذج انحدار كاما في ظل مشكلة الارتباط الخطي المتعدد. وكما هو معروف، يصبح المنهج البايزي غير فعال في النماذج التي تعاني من الارتباط الخطي المتعدد، حيث تميل المعالم المقدرة إلى أن تكون غير دقيقة. لذا، ولتحسين أداء هذا المنهج، استُخدمت إحدى تقنيات الجزاء، وهي تقنية الشبكة المرنة، لتقدير معالم نموذج انحدار كاما في ظل الارتباط الخطي المتعدد. هدفت الدراسة إلى بحث سلوك المنهج البايزي في تقدير المعالم عند تطبيق هذه التقنية، بالإضافة إلى مقارنة التقنية المقترحة بالطرق التقليدية. أظهرت النتائج تفوق وفعالية التقنية المستخدمة في عملية التقدير. أُجريت تجارب محاكاة بالإضافة إلى تحليل البيانات الحقيقية، وذلك بحساب متوسط مربع الخطأ عند مستويات ارتباط مختلفة بين المتغيرات المستقلة وفي ظل أحجام عينات متفاوتة.

الكلمات المفتاحية: Elastic net، انحدار كاما، مشكلة التعدد الخطي

Corresponding Author: E-mail: gh.gheadaa@uodiyala.edu.iq

Introduction

Regression analysis is a fundamental statistical tool used to model and explain relationships among variables. Typically, one variable is treated as the dependent variable, while the others serve as explanatory variables. These relationships are formulated through linear or nonlinear functions, and model parameters are estimated using appropriate statistical methods for interpretation and prediction.

When the dependent variable is positive and follows a Gamma distribution, as in time, biological, and engineering data, Gamma regression models within the framework of generalized linear models (GLMs) are appropriate. However, like other regression models, Gamma regression suffers from several statistical issues, most notably multicollinearity among explanatory variables, which leads to unstable estimates, inflated variances, and reduced predictive performance[5 – 8].

To address this issue, regularization methods have been developed, among which the Elastic Net technique combines the advantages of Ridge and Lasso regression to achieve both shrinkage and variable selection. In parallel, the Bayesian approach provides a flexible estimation framework by incorporating prior information and accounting for uncertainty.

Therefore, integrating the Elastic Net technique within the Bayesian framework offers an effective approach for improving parameter estimation in Gamma regression models under multicollinearity, enhancing estimation stability and predictive accuracy[1 – 3].

Research Problem

The Gamma regression model is one of the generalized linear models used when the dependent variable follows a two-parameter Gamma distribution. The first parameter represents the mean M_i , which is expressed as a linear combination of the explanatory variables through an appropriate link function, while the second parameter represents the shape parameter ϕ , which can also be modeled as a linear function of a set of covariates. Although numerous studies have examined this model and proposed various methods for estimating its parameters, the application of the Gamma regression model faces several complex statistical challenges, most notably the problem of multicollinearity among the explanatory variables[6].

3. Objective

This study aims to develop a hybrid estimation approach based on integrating the Elastic Net technique within the Bayesian framework to improve the estimation of the parameters of the Gamma regression model in the presence of multicollinearity among explanatory variables. This hybridization seeks to exploit the shrinkage and variable selection properties offered by the Elastic Net technique, while simultaneously benefiting from the probabilistic inference capabilities provided by the Bayesian approach, thereby enhancing the stability of the estimates and improving their predictive accuracy[18 – 19].

Furthermore, the study includes a comparative evaluation of the performance of the proposed hybrid Bayesian model against several other estimation methods, particularly the standard Bayesian method, the Maximum Likelihood method, and the Kibria–Lukman method[17], in order to assess its efficiency in addressing multicollinearity and improving estimation quality compared with both traditional and modern approaches.

4. GammaRegressionModel:

The Gamma regression model is considered one of the generalized linear models (GLMs) proposed by Nelder and Wedderburn (1972). This model is regarded as an extension of the general linear model, and it requires that the distribution of the response variable belongs to the exponential family. The probability density function of the Gamma distribution can be written as follows[17]:

$$f(y, \alpha, \lambda) = \frac{1}{\Gamma_{\alpha} \lambda^{\alpha}} Y^{\alpha-1} e^{-\frac{Y}{\lambda}} \quad Y, \alpha, \lambda > 0 \quad (1)$$

where α represents the shape parameter and λ represents the scale parameter. The Gamma distribution function can be reformulated in terms of the mean parameter and the shape parameter as follows:

$$E(Y) = M = \alpha \lambda$$

We assume that the shape parameter can be expressed in the following form [9].

$$\emptyset = \frac{1}{\alpha} \rightarrow \alpha = \frac{1}{\emptyset}$$

The parameter $\emptyset$ denotes the shape (or dispersion) parameter. Substituting the parameter α into the mean function yields the parameter λ , as given by:

$$M = \frac{1}{\emptyset} * \lambda$$

$$\lambda = M * \emptyset$$

Of course, the Gamma distribution function can be reformulated in terms of the two new parameters as follows[16]:

$$f(y, \emptyset, M) = \frac{1}{\Gamma_1(M\emptyset)^{\frac{1}{\emptyset}}} Y^{\frac{1}{\emptyset}-1} e^{-\frac{Y}{M\emptyset}} \quad Y \geq 0 \quad (2)$$

Following the reparameterization of the probability distribution function, the distribution can be shown to belong to the exponential family, expressed as follows[9 – 12]:

$$(y, \theta, \emptyset) = \exp \left(\frac{Y\theta - b(\theta)}{\alpha(\emptyset)} + c(y, \emptyset) \right) \quad (3)$$

Where

y : the response variable that belongs to the exponential family.

ϕ : represents the dispersion parameter (or shape parameter), which appears in distributions that involve two or more parameters.

θ : represents the location parameter or the canonical parameter.

$\alpha(\phi)$: denotes a function depending on the dispersion parameter ϕ .

$b(\theta)$: denotes a function of the canonical (location) parameter.

$c(y, \phi)$: a function of the observations y and the dispersion parameter ϕ .

The Gamma probability distribution function (2) can be re-expressed in terms of the exponential family function (3) as follows:

$$f(Y, \theta, \emptyset) = \exp \left[\frac{\frac{Y}{M} - (-\ln(\theta))}{-\emptyset} + \frac{1 - \emptyset}{\emptyset} \ln(Y) - \frac{\ln(\emptyset)}{\emptyset} - \ln \Gamma \left(\frac{1}{\emptyset} \right) \right] \quad (4)$$

where:

$$\theta = \frac{1}{M}, b(\theta) = -\ln(M), \alpha(\emptyset) = -\emptyset, c(y, \emptyset) = \frac{1 - \emptyset}{\emptyset} \ln(Y) - \frac{\ln(\emptyset)}{\emptyset} - \ln \Gamma \left(\frac{1}{\emptyset} \right)$$

The mean and variance of the Gamma distribution within the exponential family can be expressed as follows:

$$E(Y) = \frac{\partial b(\theta)}{\partial \theta} = b'(\theta) = M$$

$$\text{var}(Y) = \frac{\partial^2 b(\theta)}{\partial \theta^2} \alpha(\emptyset) = b''(\theta) \alpha(\emptyset) = \emptyset M$$

Since the Gamma distribution is considered a member of the exponential family, link functions are used to transform this distribution into a regression model. This is achieved by modifying the location parameter, which represents the mean. Although the shape parameter is often assumed to be constant in many studies, the present study assumes it to be variable. Therefore, it is estimated in a manner similar to the location parameter. Two link functions are employed: the first is an exponential link function for the mean parameter, and the second is a natural link function for the dispersion (or shape) parameter. [13]

$$\eta_{1i} = g(M_i) = e^{X' \beta} \quad (5)$$

$$\eta_{2i} = h(\emptyset_i) = Z' \gamma \quad (6)$$

$(\gamma_0, \gamma_1, \dots, \gamma_p), (\beta_0, \beta_1, \dots, \beta_p)$: These are the vectors of regression parameters corresponding to the mean and dispersion parameters

$g(M_i)$: Link function for the mean (exponential link function).

$h(\emptyset_i)$: The link function for the shape (or dispersion) parameter is the natural link function

$X_i = (X_{i1}, X_{i2}, \dots, X_{ip})'$: A vector of P independent variables.

$Z_i = (Z_{i1}, Z_{i2}, \dots, Z_{iK})'$ A vector of K independent variables.

Returning to the Gamma distribution function (2), it can be rewritten in terms of link functions to transform it into a Gamma regression model, as follows:

$$f(y, X, \gamma, \beta) = \frac{1}{\Gamma \frac{1}{Z' \gamma} (e^{X' \beta} Z' \gamma)^{\frac{1}{Z' \gamma}}} \cdot Y^{\left(\frac{1}{Z' \gamma} - 1\right)} e^{-\frac{Y_i}{Z' \gamma e^{X' \beta}}} \quad (7)$$

The above equation can be simplified to obtain the following form:

$$f(y, X, \gamma, \beta) = \frac{e^{\left[\left(-\frac{1}{Z' \gamma}\right) \log(Z' \gamma)\right]}}{\Gamma \frac{1}{Z' \gamma}} Y^{\left(\frac{1}{Z' \gamma} - 1\right)} \cdot e^{-\frac{1}{Z' \gamma} [Y_i e^{-X' \beta} + X' \beta]} \quad (8)$$

Equation (8) is considered the final form after combining the Gamma distribution with a linear regression model, resulting in a Gamma regression model. By maximizing the above function, the maximum likelihood function can be obtained:

$$L(y, X, \gamma, \beta) = \left(\frac{e^{\left[\left(-\frac{1}{Z' \gamma}\right) \log(Z' \gamma)\right]}}{\Gamma \frac{1}{Z' \gamma}} \right)^n \pi_{i=1}^n Y^{\left(\frac{1}{Z' \gamma} - 1\right)} \cdot e^{-\frac{1}{Z' \gamma} \sum [Y_i e^{-X' \beta} + X' \beta]} \quad (9)$$

5- Estimates of the Gamma Regression Model Parameters

he parameter estimates of the Gamma regression model are divided into two parts. The first part involves estimating the model parameters using Bayesian estimators and the hybrid Bayesian elastic-net technique, which is integrated within the Bayesian framework to address the multicollinearity problem in estimating the parameters of the Gamma regression model. The second part involves estimating the model parameters using the Kibria-Lukman method[11].

A- The Concept of Bayesian Theory

The concept of Bayesian theory is based on the assumption that the variables to be estimated are random variables rather than fixed values. These variables are estimated using prior information, also known as prior distributions, along with current information or observations. This information is represented by the likelihood function, and by combining these two functions, the posterior probability function is obtained[10].

In Bayesian theory, the prior distribution serves as the basis for estimating the primary variables, relying on previous studies. Prior distributions are categorized into two types: unknown priors and known priors. The posterior distribution, on the other hand, is defined as the function that incorporates all information related to the variables to be estimated. It is determined by multiplying the prior distribution of the variable by the likelihood function of the observations.

B- Using the Bayesian Approach to Estimate the Gamma Model

As previously explained in the concept of Bayesian theory, it relies on prior distributions for the parameters. Since there are two types of priors—known priors and unknown priors—this study will adopt Bayesian inference based on known prior distributions for the parameters. To assume a known distribution for the parameters, it is necessary to select a distribution that closely resembles the distribution of the dependent variable Y. Given that the dependent variable Y follows a Gamma distribution, the chosen prior for the parameters should belong to the exponential family or be a distribution defined over the positive domain. One of the most important such distributions is the exponential distribution, which is considered the closest to the Gamma distribution, as it represents a special case of the Gamma distribution. The exponential distribution function can be expressed as follows: [7]

$$(y, \lambda) = \frac{1}{\lambda} e^{-\frac{y}{\lambda}} \quad Y, \lambda > 0 \quad (10)$$

Returning to the estimates of the Gamma regression model based on Bayesian theory, which assumes the existence of a prior distribution for the parameters, it is assumed that the parameters follow an exponential distribution:

The main characteristics of the exponential distribution, represented by its mean and variance, can be expressed as follows:

Among the most important properties of the exponential distribution are the mean and the variance, which are expressed as follows:

$$m(Y) = \lambda$$

$$var(Y) = \lambda^2$$

Returning to the estimates of the Gamma regression model based on Bayesian theory, which assumes the existence of a prior distribution for the parameters, it is assumed that the parameters follow an exponential distribution:

$$\pi(\beta) = C \left(e^{-\frac{\beta_1}{\lambda_1}}, e^{-\frac{\beta_2}{\lambda_2}}, \dots, e^{-\frac{\beta_p}{\lambda_p}} \right) \quad (11)$$

where C is a diagonal matrix of order (P×P), that is:

$$C = \begin{bmatrix} \frac{1}{\lambda_1} & \dots & \dots \\ \vdots & \ddots & \vdots \\ \dots & \dots & \frac{1}{\lambda_p} \end{bmatrix}$$

β : the parameter vector ($\beta_0, \beta_1, \dots, \beta_p$).

Thus, to derive a prior distribution for the parameters, a transformation to an exponential distribution is required. This is accomplished through a link function, specifically the exponential link function, denoted by η_{3i} , which can be expressed as follows:

$$\eta_{3i} = T(M_i) = e^{(X' \beta)} \quad (12)$$

where:

$T(M_i)$ is the mean link function for the exponential distribution (the exponential link function); η_{3i} represents the linear predictors associated with the exponential distribution;

$X_i = (X_{i1}, X_{i2}, \dots, X_{ip})'$: is a P dimensional vector of independent variables

The exponential distribution function can be reformulated in the same manner as the Gamma distribution. This can be illustrated by assuming that the parameter of the exponential distribution, λ , represented by the mean, can be reparameterized so that the exponential distribution function is expressed in terms of the mean using the link function η_{3i} . This can be explained as follows:

$$\eta_{3i} = M = \lambda = e^{(X' \beta)}$$

The exponential distribution function (10) can be reparametrized in terms of the model parameters β as follows:

$$f(X, \beta) = \frac{1}{e^{(X' \beta)}} e^{-\frac{\beta}{e^{(X' \beta)}}} \quad (13)$$

Equation (13) represents the prior distribution of the model parameters β . Equation (11) can be reformulated in terms of Equation (13) as follows:

$$\prod(\beta) = C \left(e^{-\frac{\beta_1}{e^{(X_1' \beta_1)}}}, \dots, e^{-\frac{\beta_p}{e^{(X_p' \beta_p)}}} \right) \quad (14)$$

Since C is a square matrix of order $P \times P$, it can be represented in the following form:

$$C = \begin{bmatrix} \frac{1}{e^{(x_1' \beta_1)}} & \cdots & \cdots \\ \vdots & \ddots & \vdots \\ \cdots & \cdots & \frac{1}{e^{(x_p' \beta_p)}} \end{bmatrix}$$

In order to obtain the posterior distribution, according to Bayesian theory, the prior distribution of the parameters is multiplied by the likelihood function, followed by maximizing the resulting function, that is:

$$\prod (\beta/Y, X) = L(\beta/Y, X) * \prod (\beta) \quad (15)$$

This is done by multiplying the maximum likelihood function given in Equation (9) by the prior distribution equation:

$$\begin{aligned} & \prod (\beta, \gamma/Y, X) \\ &= \left(\frac{e^{[-\frac{1}{Z' \gamma}] \log(Z' \gamma)}}{\Gamma_{\frac{1}{Z' \gamma}}} \right)^n \pi_{i=1}^n Y^{\left(\frac{1}{Z' \gamma} - 1\right)} \cdot e^{-\frac{1}{Z' \gamma} \sum [y_i e^{-X' \beta + X' \beta}]} \cdot \prod \left(\frac{1}{e^{(X' \beta)}} e^{-\frac{\beta}{e^{(X' \beta)}}} \right) \end{aligned}$$

It can then be written in the following final form:

$$\begin{aligned} & \prod (\beta, \gamma/Y, X) \\ &= \left(\frac{e^{[-\frac{1}{Z' \gamma}] \log(Z' \gamma)}}{\Gamma_{\frac{1}{Z' \gamma}}} \right)^n \pi_{i=1}^n Y^{\left(\frac{1}{Z' \gamma} - 1\right)} e^{-\frac{1}{X' \beta} e^{-\frac{1}{Z' \gamma} \sum [y_i e^{-X' \beta + X' \beta}]} * \left(\frac{1}{e^{(X' \beta)}} \right)^n e^{-\sum_{i=1}^n \frac{\beta}{e^{(X' \beta)}}}} \end{aligned} \quad (16)$$

Equation (16) represents the posterior probability density function of the parameter vector β as well as the parameter vector γ . Accordingly, the hazard function can be expressed as follows:

$$h(\beta, \gamma/Y, X) = \frac{\prod (\beta, \gamma/Y, X)}{\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \prod (\beta/Y, X) d\beta_1 \cdots d\beta_p} \quad (17)$$

Equation (17) is referred to as the hazard function for the parameter vector β . A point estimate of the parameter vector can be obtained based on Bayesian theory, where the loss function, previously defined as $L=L(\hat{\beta}, \beta)$. This point estimate is determined by finding the values of $\hat{\beta}$ and $\hat{\gamma}$ that minimize the loss function. In other words:

$$\text{Min}_{\hat{\theta}} E[L(\hat{\beta}, \beta)] = \text{Min}_{\hat{\theta}} \int L(\hat{\beta}, \beta) * p(\beta/Y) d\beta \quad (18)$$

It is worth noting here that, as mentioned earlier, the squared loss function will be used to illustrate how this function operates in obtaining estimates for the regression model:

$$L(\hat{\beta}, \beta) = (\hat{\beta} - \beta)^2 \quad (19)$$

By taking the expectation of the previous equation:

$$E(L(\hat{\beta}, \beta)) = E(\hat{\beta} - \beta)^2 \quad (20)$$

After simplifying Equation (20), it takes the following form:

$$\begin{aligned} &= \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} (\hat{\beta} - \beta)^2 h(\beta/Y, X, \emptyset) d_{\beta_1} \dots d_{\beta_p} \\ &= \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} (\hat{\beta}^2 - 2\hat{\beta}\beta + \beta^2) h(\beta/Y, X, \emptyset) d_{\beta_1} \dots d_{\beta_p} \\ &= \hat{\beta}^2 \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} h(\beta/Y, X, \emptyset) d_{\beta_1} \dots d_{\beta_p} - 2\hat{\beta} \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \beta h(\beta/Y, X, \emptyset) d_{\beta_1} \dots d_{\beta_p} \\ &\quad + \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \beta^2 h(\beta/Y, X, \emptyset) d_{\beta_1} \dots d_{\beta_p} \end{aligned}$$

Since the first term is equal to one:

$$\hat{\beta}^2 \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} h(\beta/Y, X, \emptyset) d_{\beta_1} \dots d_{\beta_p} = \hat{\beta}^2$$

As for the second term, it is equal to:

$$2\hat{\beta} \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \beta h(\beta/Y, X, \emptyset) d_{\beta_1} \dots d_{\beta_p} = 2\hat{\beta} E(\beta/Y, X, \emptyset)$$

As for the third term, it is equal to:

$$\int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \beta^2 h(\beta/Y, X, \emptyset) d_{\beta_1} \dots d_{\beta_p} = E(\beta^2/Y, X, \emptyset)$$

After simplifying and evaluating the integrals for all terms, Equation (20) takes the following form:

$$E(L(\hat{\beta}, \beta)) = \hat{\beta}^2 - 2\hat{\beta} E(\beta/Y, X, \emptyset) + E(\beta^2/Y, X, \emptyset) \quad (21)$$

By differentiating $E(L(\hat{\beta}, \beta))$ with respect to $\hat{\beta}$ and equating it to zero, it follows that:

$$\frac{\partial E(L(\hat{\beta}, \beta))}{\partial \hat{\beta}} = 2\hat{\beta} - 2E(\beta/Y, X, \emptyset) = 0$$

After simplifying the function, a point estimate of the parameter vector is obtained:

$$\underline{\hat{\beta}} = E(\beta/Y, X, \emptyset) \quad (22)$$

From Equation (22), it can be seen that the point estimate of the parameter β is optimal when it coincides with the expected value of its posterior probability density function. That is, the

Bayesian estimator of β , under the squared loss function, is equal to the mean of its posterior distribution. The estimator takes the following form:

$$E(\beta/Y, X, \gamma) = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \beta h(\beta/Y, X, \Phi) d\beta_1 \dots d\beta_p \quad (23)$$

$$= \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \beta \frac{\left(\frac{e^{\left[\left(-\frac{1}{Z'\gamma} \right) \log(Z'\gamma) \right]}}{\Gamma \frac{1}{Z'\gamma}} \right)^n \pi_{i=1}^n Y^{\left(\frac{1}{Z'\gamma} - 1 \right)} e^{-\frac{1}{X'\beta}} e^{-\frac{1}{Z'\gamma} \sum [Y_i e^{-X'\beta + X'\beta}]} \left(\frac{1}{e^{(X'\beta)}} \right)^n e^{-\sum_{i=1}^n \frac{\beta}{e^{(X'\beta)}}}}{\int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \left(\frac{e^{\left[\left(-\frac{1}{Z'\gamma} \right) \log(Z'\gamma) \right]}}{\Gamma \frac{1}{Z'\gamma}} \right)^n \pi_{i=1}^n Y^{\left(\frac{1}{Z'\gamma} - 1 \right)} e^{-\frac{1}{X'\beta}} e^{-\frac{1}{Z'\gamma} \sum [Y_i e^{-X'\beta + X'\beta}]} \left(\frac{1}{e^{(X'\beta)}} \right)^n e^{-\sum_{i=1}^n \frac{\beta}{e^{(X'\beta)}}}} d\beta$$

The parameter γ is estimated using the same approach as that employed for estimating β , as described above.

$$(\gamma/Y, X, \beta) = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \gamma h(\gamma/Y, X, \Phi) d\gamma_1 \dots d\gamma_k \quad (24)$$

$$\int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \gamma \frac{\left(\frac{e^{\left[\left(-\frac{1}{Z'\gamma} \right) \log(Z'\gamma) \right]}}{\Gamma \frac{1}{Z'\gamma}} \right)^n \pi_{i=1}^n Y^{\left(\frac{1}{Z'\gamma} - 1 \right)} e^{-\frac{1}{X'\beta}} e^{-\frac{1}{Z'\gamma} \sum [Y_i e^{-X'\beta + X'\beta}]} \left(\frac{1}{e^{(X'\beta)}} \right)^n e^{-\sum_{i=1}^n \frac{\beta}{e^{(X'\beta)}}}}{\int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \left(\frac{e^{\left[\left(-\frac{1}{Z'\gamma} \right) \log(Z'\gamma) \right]}}{\Gamma \frac{1}{Z'\gamma}} \right)^n \pi_{i=1}^n Y^{\left(\frac{1}{Z'\gamma} - 1 \right)} e^{-\frac{1}{X'\beta}} e^{-\frac{1}{Z'\gamma} \sum [Y_i e^{-X'\beta + X'\beta}]} \left(\frac{1}{e^{(X'\beta)}} \right)^n e^{-\sum_{i=1}^n \frac{\beta}{e^{(X'\beta)}}}} d\gamma$$

It is evident from the above formulation that obtaining Bayesian estimates is challenging due to the presence of PPP integrals over the parameters β and γ , which complicates the computation of the Bayesian estimates. To address this issue, numerical methods are recommended for obtaining the estimates, including the Metropolis-Hastings algorithm. The Metropolis-Hastings algorithm is one of the Markov Chain Monte Carlo (MCMC) methods and is particularly useful when sampling from the target probability distributions is difficult or even impossible.

c- Elastic Net technique

In general, regularization in Gamma models plays a significant role in improving prediction accuracy. The Elastic Net technique is an important form of regularized regression, proposed by Zou & Hastie (2005) as an extension of the Lasso method, offering improved performance over Lasso[16 – 18 – 19].

In order to integrate the Bayesian approach in estimating the parameters of the Gamma regression model, after obtaining them according to the previous formulation (24) using the Metropolis-Hastings algorithm as well as the Gauss-Hermite algorithm, the Elastic Net method relies on the following formulation:

$$\min_{\beta}(\sum (Y_i - X' \beta)^2 + \delta P_{\tau}(\beta)) \quad (25)$$

Where P_{ω} represents the function of the parameters to be estimated, and it can be written as follows:

$$P_{\tau}(\beta) = \sum_{i=1}^{P_i} \left(\frac{(1-\tau)}{2} \beta_j^2 + \tau |\beta_j| \right) \quad (26)$$

Where

P_i : represents the number of features to be estimated.
 τ : denotes a probability value bounded within the interval (0,1)(0,1)(0,1).
 β_j^2 represents the **Ridge** penalty term.
 $|\beta_j|$: represents the **Lasso** penalty term.

δ : It represents the sum $\delta_1 + \delta_2$, where δ_1, δ_2 (denote the tuning and regularization parameters.

The value of τ can be obtained through the following assumption

$$\tau = \frac{\delta_2}{\delta_1 + \delta_2} \quad (27)$$

$$1 - \tau = 1 - \frac{\delta_2}{\delta_1 + \delta_2}$$

$$1 - \tau = \frac{\delta_1 + \delta_2 - \delta_2}{\delta_1 + \delta_2}$$

$$1 - \tau = \frac{\delta_1}{\delta_1 + \delta_2} \quad (28)$$

Where the Elastic Net formulation is given as follows:

$$L(\delta_1, \delta_2, \beta) = |Y - X' \beta|^2 + \delta_2 |\beta|^2 + \delta_1 |\beta| \quad (29)$$

Where:

$$|\beta|^2 = \sum_{j=1}^P \beta_j^2$$

$$|\beta| = \sum_{j=1}^P |\beta_j|$$

The Elastic Net estimator is obtained by minimizing equation (29).

$$\hat{\beta}_{elastic} = \arg \min_{\beta} \{L(\delta_1, \delta_2, \beta)\} \quad (30)$$

In order to adopt a Bayesian framework and integrate it with the Elastic Net technique, the parameters (δ_1, δ_2) , representing the regularization and tuning parameters, are assumed to be random variables and are estimated along with the model parameters β , which are assigned a prior distribution.

$$\delta_1/\beta_j \sim \exp(r_1)$$

$$\delta_2/\beta_j \sim \exp(r_2)$$

Where r_1, r_2 denote the distributional parameters associated with the random variables (δ_1, δ_2) , respectively.

Substituting into equation (26) yields $P_\tau(\beta)$, after the value of τ has been determined using equation (27).

The Elastic Net technique represents a combination of the Lasso and Ridge models. Specifically, setting $\tau=0$ results in the Ridge estimator, whereas setting $\tau=1$ results in the Lasso estimator. For implementing the Elastic Net in a Bayesian estimation framework, the following expression is employed:

$$\min_{\beta_{Bay}} \left(\sum (Y_i - X' \beta_{Bay})^2 + \delta P_\tau(\beta_{Bay}) \right) \quad (31)$$

Where:

$$P_\tau(\beta_{Bay}) = \sum_{i=1}^{P_i} \frac{(1-\tau)}{2} \beta_j^2 Bay + \tau |\beta_j Bay| \quad (32)$$

Equation (32) represents the final estimates of the Elastic Net technique and is considered one of the best estimators in the presence of multicollinearity. These estimates are obtained using the Cross-Validation method.

D- The Kibria-Lukman Method

In recent years, several researchers have proposed modified ridge regression methods that demonstrate reduced bias compared to conventional techniques. Notably, Kibria and Lukman developed a modified ridge regression approach, which has prompted comparisons with this modified method[17].

Along with several other methods in 2022, and to illustrate this approach, we proceed as follows:

$$\begin{aligned} \hat{\beta}_{GKLE} &= (X'W_1X + KI)^{-1}(X'\widehat{W}_1X - KI)\hat{\beta}_{MLE} \\ &= (F_1 + KI)(F_1 - KI)\hat{\beta}_{MLE} \\ &= D_K\hat{\beta}_{MLE} \end{aligned} \quad (33)$$

$\hat{\beta}_{MLE}$ denotes the maximum likelihood estimates, obtained via the weighted maximum likelihood algorithm.

$$D_K = (F_1 + KI)(F_1 - KI).$$

K: denotes the bias parameter.

The variance-covariance matrix of this method can be calculated as follows:

$$\begin{aligned} Cov(\hat{\beta}_{GKLE}) &= Cov(D_K \hat{\beta}_{MLE}) \\ &= \hat{\sigma}_1 D_K F_1^{-1} D_K' \end{aligned} \quad (34)$$

In order to calculate the mean squared error (MSE), the minimum mean squared error (MMSE) should be computed first.

$$\begin{aligned} MMSE(\hat{\beta}_{GKLE}) &= Cov(\hat{\beta}_{GKLE}) + B(\hat{\beta}_{GKLE})B'(\hat{\beta}_{GKLE}) \\ &= \hat{\sigma}_1 D_K F_1^{-1} D_K' + 4K^2 (F_1 + KI)^{-1} \beta \beta' ((F_1 + KI)^{-1})' \\ &= \hat{\sigma}_1 D_{K1} F_1^{-1} D_K' \end{aligned} \quad (35)$$

The estimator $\hat{\sigma}_1$ can be computed according to the following expression:

$$\hat{\sigma}_1 = (n - p)^{-1} \sum_{i=1}^n \left(\frac{Y_i - \hat{M}_l}{\hat{M}_l} \right)^2 \quad (36)$$

The mean squared error (MSE) is obtained by computing the trace of the MMSE matrix

$$\begin{aligned} MSE(\hat{\beta}_{GKLE}) &= tr(MMSE(\hat{\beta}_{GKLE})) \\ &= D_{K1} F_1^{-1} D_K' tr(D_{K1} F_1^{-1} D_K') \\ &= \hat{\sigma}_1 \sum_{i=1}^P \frac{(\varepsilon_1 - K)^2}{\varepsilon_1 (\varepsilon_1 + K)^2} + 4K^2 \sum_{i=1}^P \frac{e_{i1}^2}{(\varepsilon_1 + k)^2} \end{aligned} \quad (37)$$

It should be noted that e_{i1}^2 can be computed according to the following expression:

$$e_{i1} = A' \hat{\beta}_{MLE} \quad (38)$$

A' corresponds to the matrix of eigenvectors of F_1 , whereas ε_1 represents the eigenvalues of F_1^{-1} . Following the same methodology, the parameter γ is estimated as follows:

$$\hat{\gamma}_{GKLE} = (Z' W_2 Z + KI)^{-1} (Z' \hat{W}_2 Z - KI) \hat{\gamma}_{MLE} \quad (39)$$

Regarding the variance-covariance matrix, its final form is expressed as follows:

$$Cov(\hat{\gamma}_{GKLE}) = \hat{\sigma}_2 D_K F_2^{-1} D_K' \quad (40)$$

The final form of the mean squared error (MSE) is expressed as follows:

$$MSE(\hat{\gamma}_{GKLE}) = \hat{\sigma}_2 \sum_{j=1}^K \frac{(\varepsilon_2 - K)^2}{\varepsilon_2 (\varepsilon_2 + K)^2} + 4K^2 \sum_{j=1}^K \frac{e_{i2}^2}{(\varepsilon_2 + k)^2} \quad (41)$$

The estimator $\hat{\sigma}_2$ can be computed

$$\hat{\sigma}_2 = (n - K)^{-1} \sum_{i=1}^n \left(\frac{Y_i - \hat{\phi}_j}{\hat{\phi}_j} \right)^2 \quad (42)$$

Regarding the bias parameter K, there are several formulas used to calculate it. In this study, the Horel–Kennard formula is employed, which can be expressed as follows:

$$K_1 = \frac{\hat{\sigma}_i}{e_{i \max}^2} \quad (43)$$

Similarly, the Hoe formula is written as follows

$$K_2 = \frac{P\hat{\sigma}_i}{\hat{e}'\hat{e}} \quad (44)$$

6- Simulation

section, we explain how to construct a Gamma regression model using simulation experiments implemented in MATLAB, based on Monte Carlo experiments. The response variable $Y \sim G(M_i, \phi_j)$ is generated with two parameters: the mean parameter M_i , which is generated using the link function $M_i = \exp(X_i' \beta)$, and the shape parameter, which is generated using the link function $(\phi_i) = Z_i' \gamma$. The sample sizes are chosen as (100, 70, 50), and the correlation levels are set to (0.9, 0.95, 0.99). The function used to generate the independent variables is [1 – 3 – 15]

$$X_{ij} = (1 - \rho^2)R_{ij} + \rho R_{iq} \quad i = 1, 2, \dots, n \quad j = 1, 2, \dots, q \quad (44)$$

where ρ denotes the degree of correlation among the independent variables, and R_{ij} represents the independent variables, which follow a standard normal distribution. The parameter denotes the number of independent variables. Two models will be generated: the first is based on three independent variables, and the second on five independent variables. The model parameters are assigned assumed (preset) values for the Gamma regression model, as shown in the following table.

model	Mean Parameter				Shape parameter			
	β_0	β_1	β_2	β_3	γ_0	γ_1	γ_2	γ_3
Model1	1.8	0.1	0.02	7.2	7.1	0.2	0.1	1.9

model	Mean Parameter						Shape parameter					
	β_0	β_1	β_2	β_3	β_4	β_5	γ_0	γ_1	γ_2	γ_3	γ_4	γ_5
Model2	0.17	0.02	0.06	0.21	0.01	0.06	0.21	0.04	0.07	0.35	0.21	0.08

Regarding the number of repetitions, the experiment will be repeated 1,000 times. The mean squared error (MSE) is calculated according to the following formula:

$$MSE(\hat{\beta}) = \frac{1}{1000} \sum_{r=1}^{1000} (\hat{\beta}_r - \beta)' (\hat{\beta}_r - \beta) \quad (43)$$

After conducting the experiments based on the above information, the following results were obtained:

A. First Model

Table (1) presents the Mean Squared Error (MSE)

n	ρ	KLM		Bay	Bay _{elastic}
		K1	K2		
50	0.25	0.59	0.51	0.653	0.44
	0.95	0.529	0.502	0.737	0.42
	0.99	0.511	0.4	0.792	0.41
150	0.25	0.3876	0.356	0.56997	0.3332
	0.95	0.3493	0.348	0.5786	0.2321
	0.99	0.312	0.312	0.599	0.2152
250	0.25	0.244	0.249	0.245	0.234
	0.95	0.242	0.235	0.456	0.173
	0.99	0.221	0.221	0.496	0.123

From the above table, it can be observed that for the KLM model with two parameters (K1, K2), the mean values decrease as the correlation among the independent variables increases. Similarly, the mean squared error decreases with increasing sample size. In contrast, for the Bayesian approach, the mean squared error increases as the correlation among the independent variables rises. Although the error decreases with larger sample sizes, it remains generally higher compared to the other methods. As for the hybrid Elastic Net technique, it achieved the best mean squared error values, with lower errors observed as the correlation between variables increased and as the sample size grew, making it the most effective method in minimizing error.

The Second Model

Table (2) presents the Mean Squared Error (MSE).

n	ρ	KLM		Bay	Bay _{elastic}
		K1	K2		
50	0.25	0.921	0.944	1.998	0.845
	0.95	0.994	0.964	2.567	0.822
	0.99	1.25	1.37	2.786	0.821
150	0.25	0.813	0.837	0.996	0.723
	0.95	0.828	0.844	1.296	0.711
	0.99	0.833	0.8562	1.456	0.707
250	0.25	0.536	0.535	0.859	0.545
	0.95	0.581	0.581	0.876	0.522
	0.99	0.553	0.586	0.889	0.501

From the above tables, it can be observed that the methods are affected by the increase in the number of variables. For the KLM method, the mean squared error values increase with the number of independent variables compared to the previous model. Moreover, the values rise as the correlation among the variables increases. Regarding sample sizes, the MSE decreases with larger sample sizes.

For the Bayesian method, the approach is generally influenced by both the number of independent variables and the correlation among them, showing the highest MSE values compared to the other methods. However, the MSE decreases with increasing sample sizes.

As for the hybrid Elastic Net technique, it remains the best-performing method, achieving lower MSE values compared to the other approaches. Although it is also affected by an increase in the number of independent variables, the MSE decreases as the correlation among variables and the sample size increase

7- Practical Aspect

Since the Elastic Net method performed best in the simulation study, it was applied to real data. The data were collected from the City of Medicine Department, Hematology Center, from patients diagnosed with beta-thalassemia, a type of Mediterranean anemia. The total number of observations was 150. The dependent (response) variable in the study was bone age, measured in months.

Several independent variables believed to influence the response were selected after consultation with specialized physicians and practitioners in thalassemia. Four variables were identified as affecting the dependent variable, which also exhibited correlations among them, as follows

X1 Actual age (in months)

X2 Age at disease onset (in months)

X3 Liver enlargement (in centimeters)

X4 Blood hemoglobin

The data were tested to determine whether they follow a Gamma distribution using the statistical software SPSS. The results are presented in the following figure.

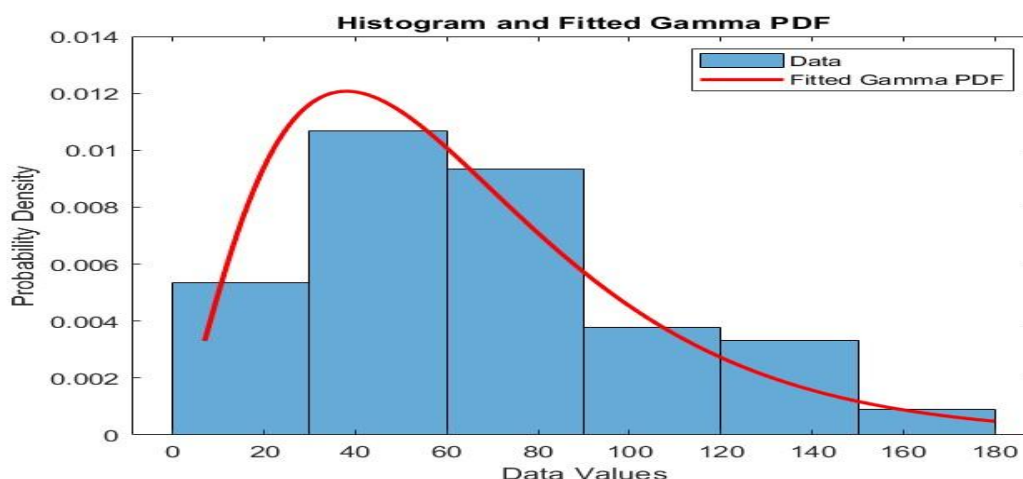


Figure (1) illustrates the probability density function of the Gamma distribution for the dependent variable (Y) based on the real data.

Chi-Square test was also conducted to assess the goodness-of-fit of the data to the Gamma regression model. The test resulted in a p-value of 0.055741, indicating that the data are suitable for the model.

After confirming that the dependent variable follows a Gamma distribution and noting that the independent variables exhibit multicollinearity, the mean of the model was calculated and compared with the simulation results. The findings are as follows:

طرق	MSE	γ_0	γ_1	γ_2	γ_3	γ_4	β_0	β_1	β_2	β_3	β_4
KLM(K1)	0.391	0.699	0.006	0.084	0.0002	0.0006	0.698	0.006	0.084	0.0002	0.0005
KLM(K2)	0.233	0.698	0.006	0.085	0.0002	0.0005	0.696	0.006	0.084	0.0002	0.0005
Bay	10.338	0.601	0.005	-0.052	0.0002	0.0007	0.595	0.005	0.05	0.0002	0.0006
<i>Bay_{elastic}</i>	0.124	0.628	0.006	-0.062	0.0001	0.0005	0.626	0.005	0.06	0.0002	0.0006

Table (3) presents the parameter estimates and the mean squared error (MSE) of the model based on the real data

From the above table, it can be observed that the hybrid Bayesian Elastic Net method outperformed the other approaches. This demonstrates the efficiency of the technique in estimating the parameters of the Gamma regression model in the presence of multicollinearity.

8- Conclusions:

Based on the simulation results, it can be concluded that the KLM method with two bias parameters is highly affected by the number of independent variables, as the error values increase with an increasing number of variables. In contrast, the Bayesian approach is inefficient in estimating the parameters of the Gamma regression model in the presence of multicollinearity, since the mean squared error (MSE) increases with higher correlations among the variables. Furthermore, this method is also significantly influenced by the number of independent variables.

On the other hand, the hybrid Elastic Net technique proves to be efficient in estimating the parameters of the Gamma regression model under multicollinearity, as the MSE decreases with increasing correlation among variables. Although it is also affected by an increase in the number of independent variables, it generally outperforms the other methods. Overall, it can be concluded that all methods are influenced by an increase in the number of independent variables due to the expanding complexity of the relationships among variables; however, the hybrid technique remains the most effective in estimating the Gamma regression model parameters in the presence of multicollinearity.

9- Recommendations:

1. It is recommended to apply the hybrid technique to datasets affected by multicollinearity for estimating the parameters of the Gamma regression model, considering it an efficient method.
2. It is recommended to hybridize other Bayesian approaches and compare them with traditional methods for estimating the Gamma regression model parameters under multicollinearity conditions.

References

1. L. A. Abdaljabbar, and Q. N. Nayef, Comparison between maximum likelihood and Bayesian methods for estimating the gamma regression with practical application, Journal of Economics and Administrative Sciences, vol. 27, no. 125, pp. xx-xx, 2021. <https://doi.org/xxxx>.
2. M. Abdulaziz Dabdoob, M. Mohammed Tayyib Al-Naemi, and Aswan, Suggested methods in ridge regression, Iraqi Journal of Statistical Sciences, vol. 6, no. 2, pp. 85-106, 2006.
3. L. A. Abdaljabbar, and Q. N. Naif, Using the shrinkage method in estimating gamma regression by simulation method, Journal of Al-Turath University College, vol. 1, no. 31, pp. 101-110, 2021.

- 4- S. F. Aboud, and I. K. Abdullah, Estimation of the extended gamma regression model with practical application, Journal of Al-Turath University College, vol. 1, no. 35, pp. 278-294, 2023.
- 5- Z. Y. Algamal, Developing a ridge estimator for the gamma regression model, Journal of Chemometrics, vol. 32, no. 10, e3054, 2018. <https://doi.org/xxxx>.
- 6- Z. Y. Algamal, Shrinkage estimators for gamma regression model, Electronic Journal of Applied Statistical Analysis, vol. 11, pp. xx-xx, 2019. <https://doi.org/xxxx>.
- 7- Z. Y. Algamal, and I. I. Allyas, Gamma regression model estimation using bootstrapping procedure, IOSR Journal of Mathematics, vol. 13, no. 3, pp. 17-23, 2017. <https://doi.org/xxxx>.
- 8- Z. Y. Algamal, and Y. Asar, Liu-type estimator for the gamma regression model, Communications in Statistics - Simulation and Computation, vol. 49, no. 8, pp. 2035-2048, 2020. <https://doi.org/xxxx>.
- 9- Z. Y. Algamal, and K. B. Rasheed, Paired bootstrapping procedure in gamma regression model using R, Journal of Basrah Researches, vol. 37, no. 4, pp. 201-211, 2011.
- 10- R. Alhamzawi, Bayesian elastic net Tobit quantile regression, Communications in Statistics - Simulation and Computation, vol. 45, no. 7, pp. 2409-2427, 2016. <https://doi.org/xxxx>.
- 11- T. Ali, A. Al-Saffar, and S. S. Ismael, Using Bayes weights to estimate parameters of a gamma regression model, Iraqi Journal of Statistical Sciences, vol. 20, no. 1, pp. 43-54, 2023.
- 12- M. C. Brown, and E. C. Cuervo, Gamma regression model with the GammaReg R package, Comunicaciones en Estadística, vol. 8, no. 2, pp. 211-223, 2015.
- 13- P. J. Brown, and J. E. Griffin, Inference with normal-gamma prior distributions in regression problems, Bayesian Analysis, vol. 5, no. 1, pp. 171-188, 2010.
- 14- R. Chattamvelli, and R. Shanmugam, Continuous distribution in engineering and the applied sciences, Morgan & Claypool, 2021.
- 15- J. W. Hardin, and J. M. Hilbe, Generalized linear models and extensions, 4th ed., Stata Press, 2018.
- 16- W. M. Hattab, A derivation of prediction intervals for gamma regression, Journal of Statistical Computation and Simulation, vol. 86, no. 17, pp. 3512-3526, 2016.
- 17- A. F. Lukman, and B. G. Kibria, Modified ridge-type estimator for the gamma regression model, Communications in Statistics - Simulation and Computation, vol. 51, no. 9, pp. 5009-5023, 2022.
- 18- H. Zou, and T. Hastie, Regularization and variable selection via the elastic net, Journal of the Royal Statistical Society Series B: Statistical Methodology, vol. 67, no. 2, pp. 301-320, 2005.
- 19- H. Zou, and H. H. Zhang, On the adaptive elastic net with a diverging number of parameters, The Annals of Statistics, vol. 37, no. 4, pp. 1733-1751, 2009. <https://doi.org/xxxx>.